UNIVERSITY OF
ILLINOIS LIBRARY
AT URBANA-CHAMPAIGN
BOOKSTACKS

Digitized by the Internet Archive
in 2011 with funding from
University of Illinois Urbana-Champaign

http://www.archive.org/details/distinguishingtw968chen

).968
p.2

BEBR

FACULTY WORKING
PAPER NO. 968

Distinguishing the Two Forms of the Constant Percentage Learning Curve Model

Joyce T. Chen
Rene P. Manes

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 968

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

July 1983

Distinguishing the Two Forms of the Constant Percentage Learning Curve Model

Joyce T. Chen, Associate Professor
Department of Accounting
University of Illinois at Chicago

Rene P. Manes, Professor
Department of Accountancy

ABSTRACT

DISTINGUISHING THE TWO FORMS OF THE CONSTANT PERCENTAGE LEARNING CURVE MODEL

The first section of the note distinguishes between two forms of the constant percentage learning curve model, i.e. the continuous and the discrete forms. In the second section, the implications for planning and for estimation of the parameters of the learning curve model are considered wherever confusion between the two forms exists.

DISTINGUISHING THE TWO FORMS OF THE
CONSTANT PERCENTAGE LEARNING CURVE MODEL

The learning curve, also variously called the progress curve, the experience curve or the improvement curve, represents a well-documented cognitive process in labor intensive assembly and other production work. Several different mathematical models have been suggested as descriptive of the phenomenon [Yelle, 1979] but only one is widely used and discussed in the managerial accounting literature, i.e., the constant-percentage or log-linear model. The log-linear model is consistent with empirical evidence, in that production learning is often fairly rapid at first but tends to level off as the limit of efficiency is approached.

What is not usually made clear in the managerial accounting literature is that there are two versions of the constant-percentage learning curve model. The one most commonly found in the literature and in textbooks is the continuous model; the other is the discrete model. Failure to distinguish between the two could lead to their misuse and to potentially serious errors of estimation. Section I of this note distinguishes between these two related but different models; section II addresses problems of computation and model parameter estimation.

I. THE CONTINUOUS AND DISCRETE MODEL

In both models, the number of hours (cost) required to perform a repetitive process (to produce the same good) decrease by a constant percentage as a function of accumulated experience. But in the continuous model, the cumulative-average time per unit is reduced by a constant percent when quantity of production is increased by a given factor, usually a doubling of production; this relationship holds for all levels of production, integer or fractional. In the discrete model, on the other hand, attention is focused on the time required for the individual units in the production series. The time of producing the new last unit, which doubles output from its previous level, is said to decrease by a constant percentage from the time of the old last unit before the start of the doubling; this relationship holds only for integer levels of the specified unit of production.

Thus the cost function hypothesized for the first version is continuous and that for the second version is discrete. As closely related as the two models are, their use is not interchangeable over all levels of output. To illustrate how confusion and error can arise when the two models are not differentiated, we first juxtapose the two versions of the learning curve in Exhibit 1 and tabulate values for each in Table 1. A graphical comparison between the two models is provided in Figure 1. The symbol X or X_* (this and other starred symbols are for the discrete learning model) denotes the ordinal cumulative number of units produced. The learning parameters adopted for both models are:

a = the number of direct labor hours required to produce the first unit of output

b = the learning exponent derived from

$\frac{\log R}{\log F}$ where R = the rate of learning or improvement represented by the constant percent decrease in hours

F = factor increase in output (usually 2)

Certain relationships between continuous average and discrete incremental unit learning models are apparent from Table 1:

Exhibit 1

Two Forms of the Constant Percentage Learning Curve Model

	The Continuous Cumulative Average Model $X \in (0, +\infty)$	The Discrete Incremental Unit Model $X_* \in [0,1,2,\ldots,\infty]$
(1) Total time of cumulative output[a]	$T = YX = aX^{b+1}$	$T_* = \sum_{n=1}^{X_*} an^b$
(2) Average time of cumulative output	$Y = aX^b$	$Y_* = T_*/X_* = \sum_{n=1}^{X_*} an^b/X_*$
(3) Incremental unit time (change in successive total times)	$I = aX^{b+1} - a(X-1)^{b+1}$	$I_* = aX_*^b$
(4) Marginal time (slope of total time)	$M = dT/dX = (b+1)aX^b$	not defined

a. Total time of the discrete model can be approximated by

$$\int_{0.5}^{X_*+0.5} an^b dn = \frac{a}{b+1} [(X_*+0.5)^{b+1} - (0.5)^{b+1}].$$

The approximated value of total time converges with the true value as X_* increases [Summers and Welsch, 1970].

Table 1

Comparison of Continuous and Discrete Learning Model Values

[for a = 100, b = log(.8)/log(2) = -.3219]

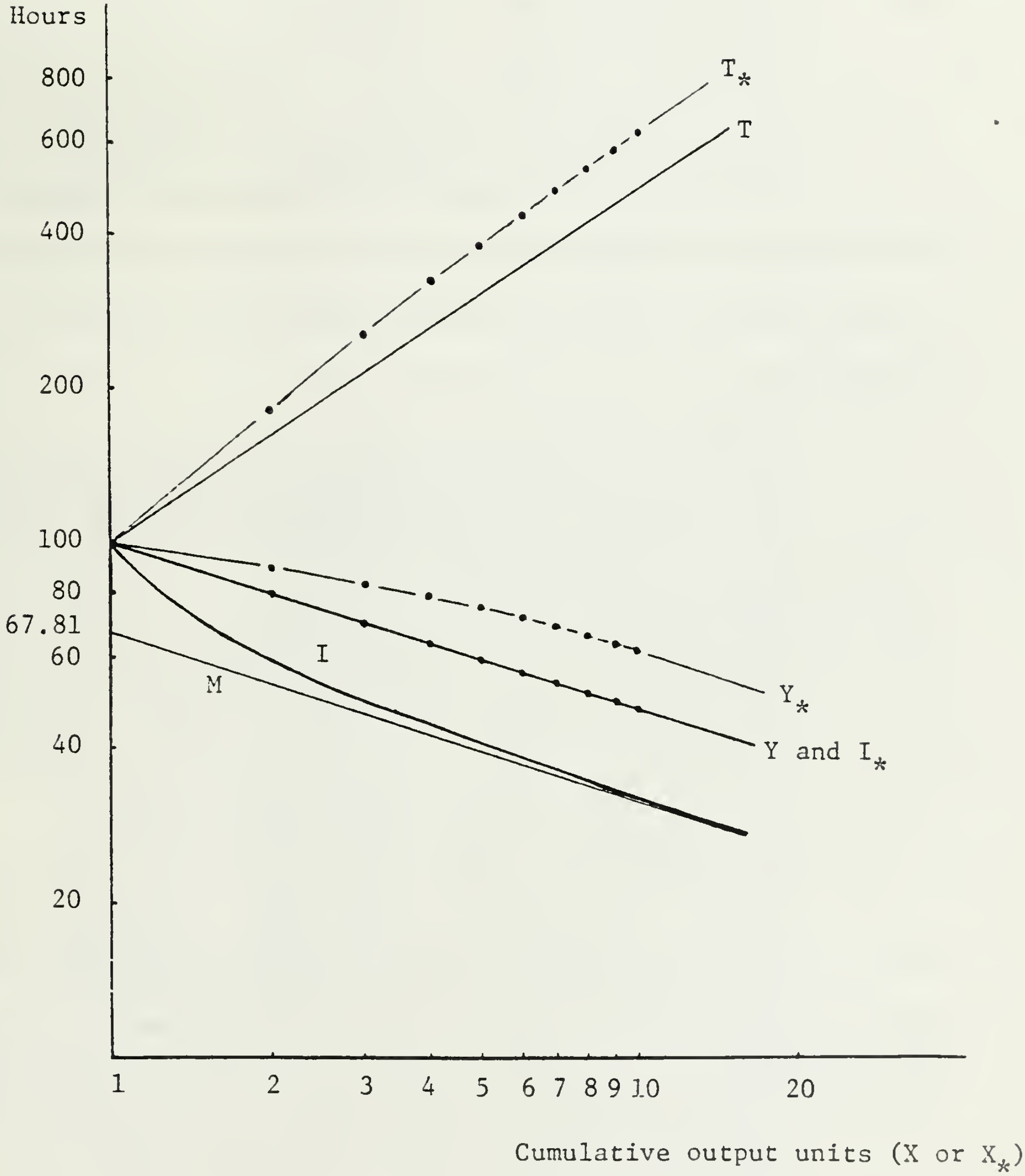
Continuous Model							Discrete Model				
Cumu. units X	Total time[a] T	Aver. time Y	Marg. time M		Incr. unit time[b] I		Cumu. units X_*	Total time T_*	Aver. time Y_*	Incr. unit time I_*	
1	100.00	100.00	67.81		100.00		1	100.00	100.00	100.00	
2	160.00	80.00	54.25	(67.81x.8)	60.00	(100x.6)	2	180.00	90.00	80.00	(100x
3	210.63	70.21	47.61		50.63		3	250.21	83.40	70.21	
4	256.00	64.00	43.40	(54.25x.8)	45.37	(60x.76)	4	314.21	78.85	64.00	(80x.
5	297.82	59.56	40.39		41.82		5	373.77	74.75	59.56	
6	337.01	56.17	38.09		39.19		6	429.94	71.65	56.17	
7	374.14	53.45	36.24		37.13		7	483.39	69.06	53.45	
8	409.60	51.20	34.72	(43.4x.8)	35.46	(45.37x.78)	8	434.59	66.83	51.20	(64x.
.	.	.	.		.			.	.	.	
.	.	.	.		.			.	.	.	
.	.	.	.		.			.	.	.	
16	655.36	40.96	27.77	(34.72x.8)	28.06	(35.46x.79)	16	892.01	55.75	40.96	(51.2

a. If the rate of learning is instead 90%, then T = (100, 180, 253.80, 324, 391.50, 457.20, 520.8, 583.20, ..., 933.12).

b. The rate of improvement in this column approaches 80% when $X \geq 128$ (see Hein 1967, p. 106

Figure 1

A Comparison of the Continuous and Discrete Learning Models Using Log-log Graph Paper



1. For the same parameter values of a and b, I* and Y have identical values, but T is consistently less than T* at all integer points after 1.

2. For the same learning parameter values, there is an identical constant percentage decline in Y, M, and I*.

3. Contrary to what may be common belief, the discrete incremental unit learning pattern is <u>not</u> identical to the continuous average learning pattern with a learning rate that is only half that of the former (that is, 1-(1-R)/2). Compare, for example, T* for an 80% learning rate with T for a 90% learning rate (see footnote a of Table 1). Although they are identical for the first two units in the production series, T* continues to diverge to lower values as cumulative production grows larger.

The second of these relationships is perhaps the main source of the confusion between the two models. Consider, for example, the following citations:

> A third representation of the learning curve can be obtained by differentiating the total-labor-hour function to obtain the <u>marginal direct labor-hour</u> (MDLH).... The marginal labor-hour function has the same exponent, b, as the average labor-hour function. Thus, the [average time] learning curve implies that both average and marginal labor-hours decrease at the learning rate, b. For example, an 80 percent learning curve is commonly found in industry....With the marginal DLH interpretation, the number of DLH for the last item produced, which doubles output from its previous level, is only 80 percent of the DLH of the item produced before we start to double output (Kaplan, 1982, p. 100).

> The learning model described so far defines Y in terms of the average number of labor hours required for X units. It is also possible to define learning in terms of the improvement in the time required to produce the <u>individual</u> unit rather than the improvement in the <u>average</u> time....Then...gives $Y' = aX^b$ where Y' is the number of labor hours required to produce the Xth unit.... Using this model the total number of hours required for n units is given by integrating [this] equation....Hence the average time for n units using this model is... $[a/(b+1)]n^b$ (Bierman-Dyckman, 1971, pp. 84-5).

Obviously a clear-cut distinction between the terms "marginal time" (slope of total time) and "incremental unit time" (change in successive total times) is essential for distinguishing the two forms of the constant percentage learning curve model.[1]

Clearly the continuous cumulative average model, which is the form most often encountered in the literature, is more tractable to deal with mathematically (although values of the discrete incremental output amounts can be easily computerized and tabulated for any values of a and b). The more important question is which model more accurately represents the learning process of a production system. Are there continuous flow processes, involving manifold labor operations and complex assembly, which are better described by the continuous cumulative model? Does the discrete incremental model better capture the learning process involving larger units and the interaction of planning and capital intensive systems with labor? [Conway and Schultz, 1959] This note does not investigate these questions other than to point out that cost data are more often collected for incremental units or batches of incremental units and that the larger the units produced, the more likely this is to be true.

II. CONSEQUENCES FOR PLANNING AND ESTIMATION

Planning

Once it is recognized that there are two distinctly separate ways of modeling the learning phenomenon, representative of two related but different cognitive processes, we must consider what the consequences are for planning and estimation.

A textbook example [Corcoran, 1978, p. 255] is used to illustrate the choices facing the firm. The problem (simplified and modified for this demonstration), is this:

> A production manager is choosing whether to make or buy a particular item. It has been estimated that it will take 80 hours to make the first unit, and that marginal time to produce unit x will be 70% of the time necessary for unit x/2. Pertinent factory costs are:
>
> | | |
> |---|---|
> | Materials | \$320 per unit |
> | Labor | \$ 5 per hour |
> | Labor related costs | \$ 3 per hour |
>
> If the item can be purchased for \$440 per unit, approximately how many units must be manufactured before the total accumulated costs under both alternatives would be equal?

In order for the manager to solve this problem, he must derive the index of learning (b = log .7/log 2 or -.5146) and formulate the indifference equation:

$$\$440x = \$320 + \$8 \text{ (total time to make a given number of units)}$$

But here the ambiguity of the term "marginal time" creates problems for the manager who must decide which specification to use for estimating the total time required to make any number of units. Should he use the formula

(i) $T = 80\, X^{1-.5146}$, or the formula

(ii) $T_* \cong \frac{80}{1-.5146}\,[(X_*+0.5)^{1-.5146} - (0.5)^{1-.5146}]$?

If he assumes the learning behavior is continuous and substitutes formula (i) into the indifference equation to yield an indifference volume, X^0 = 25.87 units and requires 388.02 hours to produce at a cost of \$11,382.40 (the same as the cost to purchase this volume externally). That is, as long as the requirement for the item is no less than 25.87

units, it is cheaper to produce in-house. On the other hand, if he assumes the discrete model is applicable and he uses formula (ii), his breakeven volume, X_*^0, = 90 units (requiring 1,350.32 hours at a cost of $39,600). Obviously, any firm which had anticipated breaking even at 26 units and thereafter saving by in-house production would be sorely disappointed to discover that its learning process would have been better described by the discrete model. The potential for unhappiness is graphically demonstrated by Figure 2.

Estimation Parameters

Thus, if there are two separate models available for characterizing the learning phenomenon, a decision must be made as to which is more representative of the actual process at work in a given case. That choice may be made on the basis of prior experience with the specific production process or by inference made by comparing estimates of the parameters common to the two models.

Suppose a company has started production of a piece of equipment subject to the learning process and has carefully accumulated data on the hours required to produce the first 8 units (refer to Table 1). Both Y and I_* are log linear, whereas I and Y_* are not. In order to determine whether the continuous or the discrete model applies, perform a logarithmic linear regression first of the cumulative average times with output and then of the incremental unit times with output; that is,

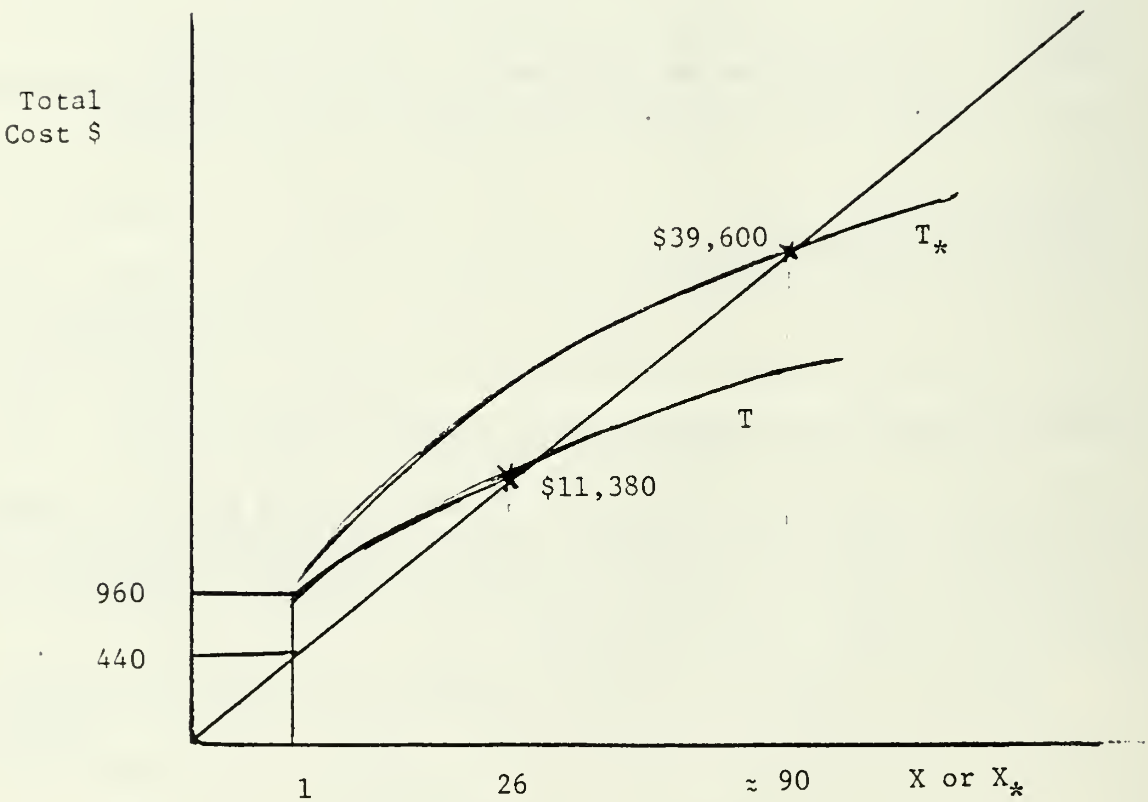
$$\log\ (\widehat{\text{time, average or incremental unit}})$$

$$= \log \hat{a} + \hat{b} \log\ (\text{cumulative units}).$$

Figure 2

Comparison of Breakeven Costs for

T and T_* with External Purchases



[not to scale]

If R^2 is higher for the logarithmic average series, the appropriate model is the continuous cumulative average model (compare $\hat{Y}$ and $\hat{I}$ in Table 2). But if R^2 is higher for the incremental than for the average series, the discrete incremental unit model is more representative of the data (compare $\hat{Y}_*$ and $\hat{I}_*$ in Table 2).

While the fit--as measured by high R^2--for the misspecified models (refer to $\hat{I}$ and $\hat{Y}_*$ in Table 2) seems quite good, a visual inspection of the scatter diagram or a sequential plot of the logarithmic residuals (Figure 3) reveals a high dependence among successive residuals. Specifically, an examination of the residuals for the misspecified continuous model reveals positive residuals for low and high values of X and negative residuals for middle values of X.[2] On the other hand, a pattern of negative-positive-negative in the residuals is observed for the misspecified discrete model.[3] Both patterns of residuals are strong evidence of an underlying logarithmic curvilinear relationship in I and Y_*.[4]

A final estimation problem is one suggested by Kaplan in his new text [1982, pp. 104-5]. He poses a problem in which the costs of the first ten units are unknown but in which incremental costs are given for units 11 to 19. Kaplan suggests estimating a logarithmic linear regression on the incremental unit time data (i) to approximate the marginal time function (M):

$$\log \hat{M} = \log (\hat{b}+1)\hat{a} + \hat{b} \log X, \text{ which implies}$$

$$\log \hat{Y} = \log \hat{a} + \hat{b} \log X.$$

Table 2

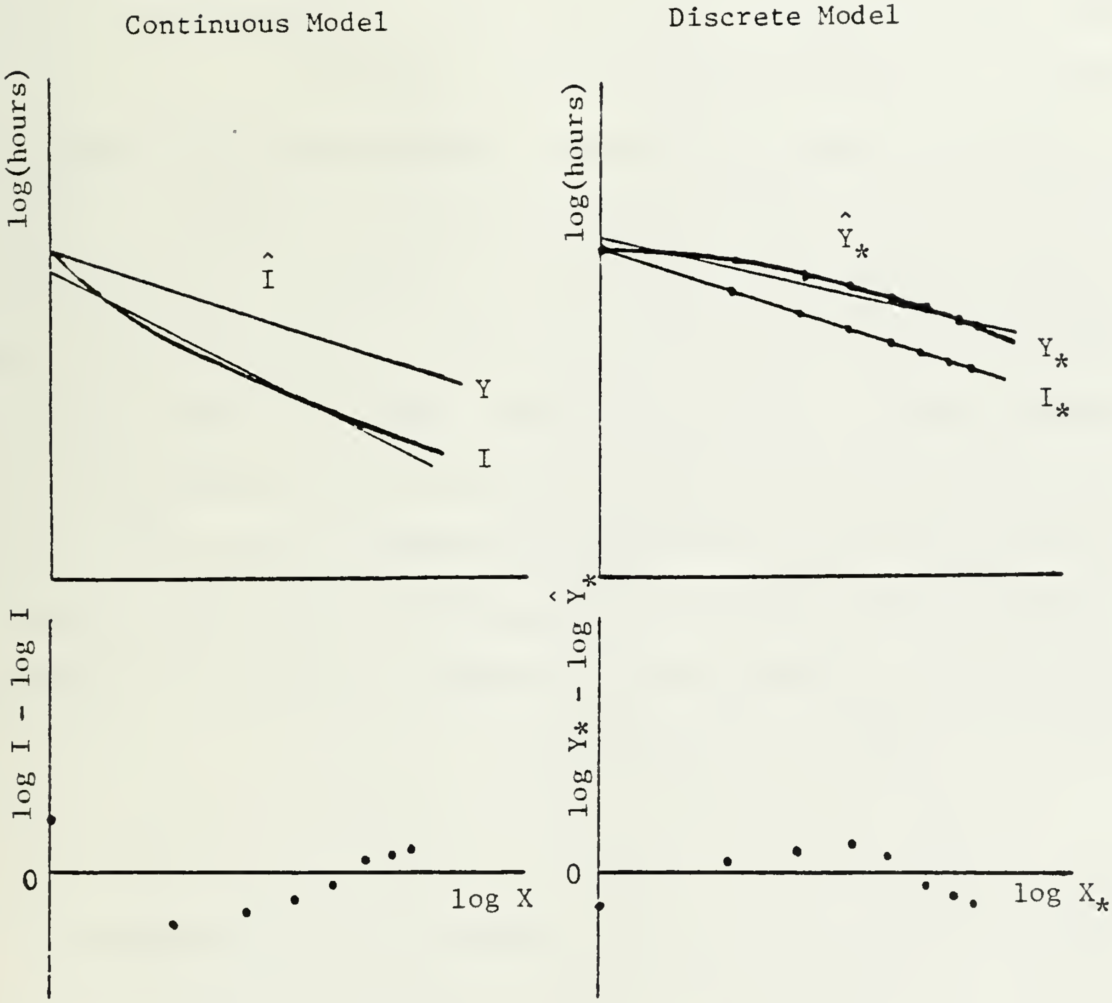
The Logarithmic Linear Regression Results[a]

	Continuous Model		Discrete Model	
	$\hat{Y}$	$\hat{I}$	$\hat{Y}_*$	$\hat{I}_*$
$\hat{a}$	100	90.91	101.95	100
$\hat{b}$	-.3219	-.4766	-.2130	-.3219
R^2	1.00	0.97	0.99	1.00

a. Based on the set of data prior to the point X (or X_*) = 8 in Table 1.

Figure 3

Patterns of Sequential Residuals for the Misspecified Learning-Curve Regression Analyses



Due to the asymptotic nature of I--that is $I \doteq M$ when X is large--Kaplan's procedure can be a valid and practical approach to approximating the continuous average time learning model. The word "asymptotic," however, means that the constant percentage learning rate in M (and also in Y) will be closely approximated but never quite reached by I. Thus, Kaplan's procedure can be valid only if substantial data from the start of the process is missing from the record.[5]

There is, however, still another, different, but quite closely related difficulty with this estimation problem. This has to do with the information loss associated with the missing data. Consider that Kaplan's problem is posed as if it is known *a priori* that a continuous learning model, as opposed to a discrete model, is at work. But, due to the substantial loss of data at the start of the process and due to the asymptotic nature of the difference between I and M, there is no way of knowing whether Kaplan's formulation is the correct one to approximate M for the continuous model, or whether we have in fact estimated I_* for the discrete model expressed as

$$\log \hat{I}_* = \log \hat{a} + \hat{b} \log X_*.$$

For ensuing production, using the existing facilities and work force, it may not matter which model had been estimated. But for a new plant, with new employees in a new location, it will make a significant difference in early start-up costs.

CONCLUSION

This note has discussed the difference between two forms of the constant percentage learning curve model. Consequences of misspecifying the learning curve process together with problems of estimating parameters of the models have also been considered. The applicability of either of the models to various production processes is an area for further investigation.

FOOTNOTES

[1]There is little advantage to be gained from cataloguing all of the accounting articles and textbooks which either lack precision or simply confuse the two models. A few other recent examples, listed alphabetically, which transgress in this regard are Belkaoui [1983], Cashin and Polimeni [1981], Copeland and Sullivan [1977], and Corcoran [1978]. See also Howell and Teichroew [1963] for the same deficiency in the quantitative methods literature.

[2]This series of logarithmic residuals is (0.0414, -0.0370, -0.0268, -0.0147, -0.0041, 0.0055, 0.0139, 0.0216).

[3]This series of logarithmic residuals is (-0.0084, 0.0048, 0.0062, 0.0063, 0.0021, -0.0008, -0.0037, -0.0066).

[4]In testing for these serial correlations, visual inspection may be advisable. The Durbin Watson test may be inappropriate for several reasons: first, the values of residuals are not homogeneous, those of early output being much larger than subsequent ones; second, confidence tables of the Durbin Watson test typically do not report intervals for less than 15 observations. If the incremental model applies to large units of output (tanks, space shuttles, submarines), readily available tables might never be applicable.

[5]Note that unless the missing data refers to a substantial amount of early output, a logarithmic regression on the incremental unit time data is theoretically incorrect for estimating the continuous learning parameters. It is to be used, rather, for estimating the discrete learning parameters. To estimate the former for cases in which the data for the first few units, say two, are missing from the record, one may consider the following relationship:

$$T(X) - T(2) = a\,(X^{b+1} - 2^{b+1}).$$

Inputs to this regression will be the cumulative incremental time required to produce beyond the second unit. Notice, however, that this regression function can no longer be linearized by a logarithmic transformation. The result is thus a more difficult to solve nonlinear problem.

REFERENCES

Belkaoui, Ahmed, Cost Accounting: A Multidimensional Emphasis (The Dryden Press 1983), pp. 287-307.

Bierman, Harold, Jr., and T. R. Dyckman, Managerial Cost Accounting (Macmillan 1971), pp. 82-89.

Buckley, John W., and Kevin M. Lightner, Accounting: An Information Systems Approach (Dickenson 1973), pp. 1087-1114.

Cashin, James A., and Ralph S. Polimeni, Cost Accounting (McGraw Hill 1981), pp. 98-99.

Conway, R. W., and A. Schultz, Jr., "The Manufacturing Process Function," The Journal of Industrial Engineering (January-February 1959), pp. 39-53.

Copeland, Benny R., and N. G. Sullivan, Cost Accounting: Accumulation, Analysis and Control (West Publishing Co. 1977), pp. 476-86.

Corcoran, Wayne, Costs: Accounting, Analysis, and Control (Wiley 1978), pp. 249-56.

Harvey, David W., "Financial Planning Information for Production Start-ups," The Accounting Review (October 1976), pp. 838-45.

Hein, Leonard W., The Quantitative Approach to Managerial Decision (Prentice-Hall 1967), pp. 90-113.

Howell, James E., and D. Teichroew, Mathematical Analysis for Business Decisions (Irwin 1963), pp. 158-63.

Kaplan, Robert S., Advanced Management Accounting (Prentice-Hall 1982), pp. 104-5.

Summers, Edward L., and Glenn A. Welsch, "How Learning Curve Models can be Applied to Profit Planning," Management Services (March-April 1970), pp. 45-50.

Yelle, L., "The Learning Curve: Historical Review and Comprehensive Survey," Decision Sciences (April 1979), pp. 302-328.

HECKMAN
BINDERY INC.
JUN 95
Bound -To -Please® N. MANCHESTER,
INDIANA 46962

UNIVERSITY OF ILLINOIS-URBANA
3 0112 060296131